GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:57:35 ; Search time 19 Seconds
(without alignments)
1289.963 Million cell updates/sec

Title: US-10-027-000-2
Perfect score: 4391
Sequence: 1 MADIDVEAILKKLTLAEKVD.....DGVALRGKFTVGETYWWSGV 833

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	969	22.1	769	3	US-09-320-878-12	Sequence 12, Appl
2	953	21.7	721	4	US-09-134-078-19	Sequence 19, Appl
3	951.5	21.7	3782	4	US-09-105-537-4	Sequence 4, Appli
4	944.5	21.5	809	4	US-09-105-537-24	Sequence 24, Appl
5	860	19.6	735	4	US-09-147-236-7	Sequence 7, Appli
6	731.5	16.7	744	2	US-08-462-080B-2	Sequence 2, Appli
7	731.5	16.7	744	3	US-08-462-090-2	Sequence 2, Appli
8	731.5	16.7	744	3	US-08-463-461-2	Sequence 2, Appli
9	433.5	9.9	804	4	US-08-981-446B-3	Sequence 3, Appli
10	132.5	3.0	328	1	US-08-386-727-6	Sequence 6, Appli
11	132.5	3.0	328	2	US-08-600-452A-6	Sequence 6, Appli
12	128.5	2.9	2032	4	US-09-071-035-458	Sequence 458, App
13	128.5	2.9	2032	4	US-09-071-035-462	Sequence 462, App
14	128.5	2.9	2032	4	US-09-071-035-466	Sequence 466, App
15	122.5	2.8	737	4	US-09-071-035-460	Sequence 460, App
16	121.5	2.8	2647	2	US-08-583-562B-8	Sequence 8, Appli
17	121.5	2.8	2647	2	US-08-779-113-8	Sequence 8, Appli
18	121	2.8	2353	4	US-09-377-155-33	Sequence 33, Appl
19	121	2.8	2353	4	US-08-913-942-4	Sequence 4, Appli
20	121	2.8	2353	4	US-09-669-974-33	Sequence 33, Appl
21	121	2.8	2411	4	US-09-268-347-36	Sequence 36, Appl
22	118.5	2.7	599	1	US-08-172-331B-4	Sequence 4, Appli
23	118.5	2.7	599	2	US-09-032-315-6	Sequence 6, Appli
24	118.5	2.7	599	2	US-08-993-318A-6	Sequence 6, Appli
25	118.5	2.7	599	4	US-09-399-886-6	Sequence 6, Appli
26	118.5	2.7	599	4	US-09-396-260-6	Sequence 6, Appli
27	118.5	2.7	599	4	US-09-576-281-6	Sequence 6, Appli
28	115	2.6	1938	4	US-09-514-302-2	Sequence 2, Appli
29	113.5	2.6	833	4	US-09-514-302-3	Sequence 3, Appli
30	113.5	2.6	2354	4	US-09-268-347-47	Sequence 47, Appl
31	111.5	2.5	1394	4	US-08-296-791-2	Sequence 2, Appli
32	111.5	2.5	1394	5	PCT-US95-10661A-2	Sequence 2, Appli
33	109.5	2.5	540	1	US-08-367-227-2	Sequence 2, Appli
34	108.5	2.5	1912	1	US-08-409-995-4	Sequence 4, Appli
35	108.5	2.5	1912	3	US-08-685-467-4	Sequence 4, Appli
36	107.5	2.4	521	4	US-08-952-365-4	Sequence 4, Appli
37	107	2.4	928	4	US-09-514-599-4	Sequence 4, Appli
38	106.5	2.4	1751	4	US-09-136-574A-44	Sequence 44, Appl
39	104.5	2.4	943	4	US-09-397-885-5	Sequence 5, Appli
40	104.5	2.4	1037	4	US-09-134-001C-4794	Sequence 4794, Ap
41	103.5	2.4	540	4	US-08-952-365-6	Sequence 6, Appli
42	103	2.3	928	1	US-08-474-140-11	Sequence 11, Appl
43	103	2.3	928	1	US-08-477-630-11	Sequence 11, Appl
44	103	2.3	928	1	US-08-472-293-11	Sequence 11, Appl
45	103	2.3	928	1	US-08-474-545-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-320-878-12
; Sequence 12, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 769
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-12

Query Match 22.1%; Score 969; DB 3; Length 769;
Best Local Similarity 32.6%; Pred. No. 1.4e-86;
Matches 277; Conservative 112; Mismatches 340; Indels 120; Gaps 25;

```
Qy      13 LTLAEKVDLLAGIDFW----------HTKALPKHGVPSLRFTDGPNGVRGTKFFNGVPAA 62
           :|| ||:  :     |          :   :|: |:| ||  |||||:|      | |
Db       1 MTLDEKISFV----HWALDPDRQNVGYLPGVPRLGIPELRAADGPNGIR----LVGQTAT 52

Qy      63 CFPCGTSLGSTFNQTLLEEAGKMMGKEAIAKSAHVILGPTINMQRSPLGGRGFESIGEDP 122
             |   :| |||: |: :  ||:||::  | :  ::||| :|  | | ||| ;|:  |||
Db      53 ALPAPVALASTFDDTMADSYGKVMGRDGRALNQDMVLGPMMNNIRVPHGGRNYETFSEDP 112

Qy     123 FLAGLGAAALIRGIQSTGVQATIKHFLCNDQEDRRMMVQSIVTERALREIYALPFQIAVR 182
            ::   | | |:|||  |:  | |||  |:||: |  | : | |: |||    |: |
Db     113 LVSSRTAVAQIKGIQGAGLMTTAKHFAANNQENNRFSVNANVDEQTLREIEFPAFE-ASS 171
```

```
Qy    183 DSQPGAFMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMSDWYGTYSTTEAVVAGLDL 242
          :  |:|| ||||:||     | : |: :|| :||: | :|||| |   |:|:  |||
Db    172 KAGAGSFMCAYNGLNGKPSCGNDELLNNVLRTQWGFQGWVMSDWLAT-PGTDAITKGLDQ 230

Qy    243 EM-----------PGPP-RFRGETLKFNVSNGK-PFIHVIDQRAREVLQFVKKCAASGV 288
          ||           | || :| || ||  | ||  |   |     | | |  |
Db    231 EMGVELPGDVPKGEPSPPAKFFGEALKTAVLNGTVPEAAVTRSAERIVGQMEKFGLLLAT 290

Qy    289 TENGPETTVNNTPETAALLRKVGNEGIVLLKNENNVLPLS--KKKKTLIVGPNAKQATYH 346
              ||    :     |: |||   | |||:||   |||:    |   ::|| |
Db    291 PAPRPE---RDKAGAQAVSRKVAENGAVLLRNEGQALPLAGDAGKSIAVIGPTAVDPKVT 347

Qy    347 GGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQCLTPDGAPGMRWRVFN 406
          | ||| :    |  |. | :  :          ||  ||    ||:          |    ||
Db    348 GLGSAHVVPDSAAAPLDTIKAR--------AGAGATVTYETGEETFGTQIPAGNLSPAFN 399

Qy    407 EPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELGLVVCGT 466
          :                  | ::   |  |   |   :|| |   |  | : :   |
Db    400 QG------------------HQLE---PGKAGALY---DGTLTVPADGEYRIAVRATG- 433

Qy    467 AKAYVDDQLVVDNATKQVPGDAFFGSATREETGRINLVKGNTYKFKIEFGSAPTYTLKGD 526
              |   ||     :  :     :|  :   :  : | || |:|     .       |: |
Db    434 --GYATVQL----GSHTIEAGQVYGKVS---SPLLKLTKG-THKL----------TISGF 473

Qy    527 TIVPGHGSLRVGGCKVIDDQAEIEKSVALAKEHDQVIICAGLNADWETEGADRASMKLPG 586
           :    || :|     | | |:|  |::    :: |    |  ||| || :: |||
Db    474 AMSATPLSLELGWVTPAAADATIAKAVESARKARTAVVFA---YDDGTEGVDRPNLSLPG 530

Qy    587 VLDQLIADVAAANPNTVVVMQTGTPEEMPWLDATPAVIQAWYGGNETGNSIADVVFGDYN 646
            |:||: || |||||:||: ||:   ||||  | ||:  || |     : | :::|| |
Db    531 TQDKLISAVADANPNTIVVLNTGSSVLMPWLSKTRAVLDMWYPGQAGAEATAALLYGDVN 590

Qy    647 PSGKLSLSFPKRLQDNPAFLNFRTEAG----------RTLYGEDVYVGYRYYEFADKDVN 696
          |||||: ||       ||  |    ||          :  | | ::|||:::  :
Db    591 PSGKLTQSF-------PAAENQHAVAGDPTSYPGVDNQQTYREGIHVGYRWFDKENVKPL 643

Qy    697 FPFGHGLSYTTFAFSNLSVSH-KDGKLSVSLSVKNTGSVPGAQVAQLYVKPLQAAKINRP 755
          ||||||||||:|  |  :|     | | |:::|:|:|   | :| | |:       :
Db    644 FPFGHGLSYTSFTQSAPTVVRTSTGGLKVTVTVRNSGKRAGQEVVQAYLGASPNVTAPQA 703

Qy    756 VKELKGFAKVELQPGETKAVTIEEQEKYVAAYFDEERDQWCVEKGDYEVIVSDSSAAKDG 815
          |:| |: || |  || | ||:    : :  ::|   | |    |:  :    |||
Db    704 KKKLVGYTKVSLAAGEAKTVTVNVDRRQL-QFWDAATDNWKTGTGNRLLQTGSSSA---- 758

Qy    816 VALRGKFTV 824
          .    |||  ||
Db    759 -DLRGSATV 766


RESULT 2
US-09-134-078-19
; Sequence 19, Application US/09134078
; Patent No. 6368844
;  GENERAL INFORMATION:
;    APPLICANT:  Bylina, Edward J.
;    TITLE OF INVENTION:  GLYCOSIDASE ENZYMES
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Gray Cary Ware & Freidenrich LLP
;      STREET:  4365 Executive Drive, Suite 1600
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/134,078
;      FILING DATE:  13-AUG-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/949,026
;      FILING DATE:  10-OCT-1997
;      APPLICATION NUMBER:  60/056,916
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  09010/024002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  858/677-1456
;      TELEFAX:  858/677-1465
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  721 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-134-078-19

  Query Match             21.7%;  Score 953;  DB 4;  Length 721;
  Best Local Similarity   30.3%;  Pred. No. 4.6e-85;
  Matches  261;  Conservative  127;  Mismatches  286;  Indels  186;  Gaps   25;

Qy      6 VEAILKKLTLAEKVDLLAGIDF--------------WHTKALPKHGVPSLRFTDGPNGV 50
            :: || :||  ||| |: |:                  |  :|: |:|:    ||| |:
Db      4 IDEILSQLTTEEKVKLVVGVGLPGLFGNPHSRVAGAAGETHPVPRLGIPAFVLADGPAGL 63

Qy     51 R----GTKFFNGVPAACFPCGTSLGSTFNQTLLEEAGKMMGKEAIAKSAHVILGPTINMQ 106
          |          |       ||    | ||:|: |||| || ||:|       |:| | :|:
Db     64 RINPTRENDENTYYTTAFPVEIMLASTWNRDLLEEVGKAMGEEVREYGVDVLLAPAMNIH 123

Qy    107 RSPLGGRGFESIGEDPFLAGLGAAALIRGIQSTGVQATIKHFLCNDQEDRRMMVQSIVTE 166
          |:|| || ||   ||| |:|  |:| ::|:|| || | |||| |:||  ||:| :||:|
Db    124 RNPLCGRNFEYYSEDPVLSGEMASAFVKGVQSQGVGACIKHFVANNQETNRMVVDTIVSE 183

Qy    167 RALREIYALPFQIAVRDSQPGAFMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMSDW 226
          |||||||   |:|||: ::|   |:||| :||  ||:|   |  :||:|||: | :||||
Db    184 RALREIYLKGFEIAVKKARPWTVMSAYNKLNGKYCSQNEWLLKKVLREEWGFGGFVMSDW 243

Qy    227 YGTYSTTEAVVAGLDLEMPGPP-------RFRGETLKFNVSNGKPFIHVIDQRAREVLQF 279
          |   :  | : || |: |||          |   | :   :  ||    |:|:  | :|:
Db    244 YAGDNPVEQLKAGNDMIMPGKAYQVNTERRDEIEEIMEALKEGKLSEEVLLECVRNILK- 302

Qy    280 VKKCAASGVTENGPE---TTVNNTPET---AALLRKVGNEGIVLLKNENNVLPLSKKKKT 333
                |  | |       :| |:    | :  : | ||:|||:| | !||  :
Db    303 --------VLVNAPSFKGYRYSNKPDLESHAEVAYEAGAEGVVLLEN-NGVLPFDENTHV 353

Qy    334 LIVGPNAKQATYHGGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQCLT 393
           : |    :    | ||      | ::  :|: :
Db    354 AVFGTGQIETIKGGTGSGDTHPRYTISILEGIKE------------------------- 387

Qy    394 PDGAPGMRWRVFNEPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADE 453
                |  |:|   :   ::| :       |   : | |:  |:|
Db    388 -------RNMKFDEELASTYEEYIKK-------MRETEEYKPR-TDSW------------ 420

Qy    454 DCTYELGLVVCGTAKAYVDDQLVVDNATKQVPGDAFFGSATREETGRINLVKGNTYKFKI 513
              | |:    |  : :  : :   |: |               |: : |  |  |
Db    421 ------GTVI----KPKLPENFLSEKEIKKPPKKNDVAVVV------ISRISGEGYDRK- 463

Qy    514 EFGSAPTYTLKGDTIVPGHGSLRVGGCKVIDDQAEIEKSVALAKE-HDQVIICAGLNADW 572
                  :|||  :              ||: |: |:|  :|| |||
Db    464 --------PVKGDFYLS------------DDELELIKTV--SKEFHDQ----------- 489

Qy    573 ETEGADRASMKLPGVLDQLIADVAAANPNTVVVMQTGTPEEM-PWLDATPAVIQAWYGGN 631
                                       ||::  |:| |:  | |    ::  |  |
Db    490 -------------------------GKKVVVLLNIGSPIEVASWRDLVDGILLVWQAGQ 523

Qy    632 ETGNSIADVVFGDYNPSGKLSLSFPKRLQDNPAFL---NFRTEAGRTLYGEDVYVGYRYY 688
          | |  :|||: |  ||||||  :|||   | |::      :    | :| ||:|||||
Db    524 EMGRIVADVLVGKINPSGKLPTTFPKDYSDVPSWTFPGEPKDNPQRVVYEEDIYVGYRYY 583
```

```
Qy    689 EFADKDVNFPFGHGLSYTTFAFSNLSVSHKDGKLSVSLSVKNTGSVPGAQVAQLYVKPLQ 748
Db    584 DTFGVEPAYEFGYGLSYTKFEYKDLKIAIDGETLRVSYTITNTGDRAGKEVSQVYIK-AP 642

Qy    749 AAKINRPVKELKGFAKVE-LQPGETKAVTIEEQEKYVAAYFDEERDQWCVEKGDYEVIVS 807
Db    643 KGKIDKPFQELKAFHKTKLLNPGESEEISLEIPLRDLASFDGKE---WVVESGEYEVRVG 699

Qy    808 DSSAAKDGVALRGKFTV-GE 826
Db    700 ASSR---DIRLRDIFLVEGE 716
```

RESULT 3
US-09-105-537-4
; Sequence 4, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3782
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-4

Query Match 21.7%; Score 951.5; DB 4; Length 3782;
Best Local Similarity 31.4%; Pred. No. 1.3e-83;
Matches 275; Conservative 123; Mismatches 340; Indels 137; Gaps 25;

```
Qy      9 ILKKLTLAEKVDLLAGIDFW----------HTKALPKHGVPSLRFTDGPNGVRGTKFFNG 58
Db   1121 LVAQMTLDEKISFV----HWALDPDRQNVGYLPGVPRLGIPELRAADGPNGIR----LVG 1172

Qy     59 VPAACFPCGTSLGSTFNQTLLEEAGKMMGKEAIAKSAHVILGPTINMQRSPLGGRGFESI 118
Db   1173 QTATALPAPVALASTFDDTMADSYGKVMGRDGRALNQDMVLGPMMNNIRVPHGGRNYETF 1232

Qy    119 GEDPFLAGLGAAALIRGIQSTGVQATIKHFLCNDQEDRRMMVQSIVTERALREIYALPFQ 178
Db   1233 SEDPLVSSRTAVAQIKGIQGAGLMTTAKHFAANNQENNRFSVNANVDEQTLREIEFPAFE 1292

Qy    179 IAVRDSQPGA--FMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMSDWYGTYSTTEAV 236
Db   1293 AS---SKAGAASFMCAYNGLNGKPSCGNDELLNNVLRTQWGFQGWVMSDWLAT-PGTDAI 1348

Qy    237 VAGLDLEM----------PGPP-RFRGETLKFNVSNGK-PFIHVIDQRAREVLQFVKK 282
Db   1349 TKGLDQEMGVELPGDVPKGEPSPPAKFFGEALKTAVLNGTVPEAAVTRSAERIVGQMEKF 1408

Qy    283 CAASGVTENGPETTVNNTPETAALLRKVGNEGIVLLKNENNVLPLS--KKKKTLIVGPNA 340
Db   1409 GLLLATPAPRPE---RDKAGAQAVSRKVAENGAVLLRNEGQALPLAGDAGKSIAVIGPTA 1465

Qy    341 KQATYHGGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQCLTPDGAPGM 400
Db   1466 VDPKVTGLGSAHVVPDSAAAPLDTIKAR--------AGAGATVTYETGEETFGTQIPAGN 1517

Qy    401 RWRVFNEPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELG 460
Db   1518 LSPAFNQG----------HQLE---PGKAGALY---DGTLTVPADGEYRIA 1552

Qy    461 LVVCGTAKAYVDDQLVVDNATKQVPGDAFFGSATREETGRINLVKGNTYKFKIEFGSAPT 520
Db   1553 VRATG---GYATVQL----GSHTIEAGQVYGKVS---SPLLKLTKG-THKL--------- 1592
```

```
Qy    521 YTLKGDTIVPGHGSLRVGGCKVIDDQAEIEKSVALAKEHDQVIICAGLNADWETEGADRA 580
Db   1593 -TISGFAMSATPLSLELGWVTPAAADATIAKAVESARKARTAVVFA---YDDGTEGVDRP 1648

Qy    581 SMKLPGVLDQLIADVAAANPNTVVVMQTGTPEEMPWLDATPAVIQAWYGGNETGNSIADV 640
Db   1649 NLSLPGTQDKLISAVADANPNTIVVVLNTGSSVLMPWLSKTRAVLDMWYPGQAGAEATAAL 1708

Qy    641 VFGDYNPSGKLSLSFPKRLQDNPAFLNFRTEAG----------RTLYGEDVYVGYRYYEF 690
Db   1709 LYGDVNPSGKLTQSF-------PAAENQHAVAGDPTSYPGVDNQQTYREGIHVGYRWFDK 1761

Qy    691 ADKDVNFPFGHGLSYTTFAFSNLSVSH-KDGKLSVSLSVKNTGSVPGAQVAQLYVKPLQA 749
Db   1762 ENVKPLFPFGHGLSYTSFTQSAPTVVRTSTGGLKVTVTVRNSGKRAGQEVVQAYLGASPN 1821

Qy    750 AKINRPVKELKGFAKVELQPGETKAVTIEEQEK----------YVAAYFDE 790
Db   1822 VTAPQAKKKLVGYTKVSLAAGEAKTVTVNVDRRQLQTGSSSADLRGSATVNVWMSSRAET 1881

Qy    791 ERDQWCVEKGDYEVIVSDSSAAKDGVALRGKFTVG 825
Db   1882 PRVPFLDLKAAYEELRAETDAAIARVLDSGRYLLG 1916
```

RESULT 4
US-09-105-537-24
; Sequence 24, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 809
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-24

Query Match 21.5%; Score 944.5; DB 4; Length 809;
Best Local Similarity 32.2%; Pred. No. 4e-84;
Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;

```
Qy      9 ILKKLTLAEKVDLLAGIDFW----------HTKALPKHGVPSLRFTDGPNGVRGTKFFNG 58
Db     56 LVAQMTLDEKISFV----HWALDPDRQNVGYLPGVPRLGIPELRAADGPNGIR----LVG 107

Qy     59 VPAACFPCGTSLGSTFNQTLLEEAGKMMGKEAIAKSAHVILGPTINMQRSPLGGRGFESI 118
Db    108 QTATALPAPVALASTFDDTMADSYGKVMGRDGRALNQDMVLGPMMNNIRVPHGGRNYETF 167

Qy    119 GEDPFLAGLGAAALIRGIQSTGVQATIKHFLCNDQEDRRMMVQSIVTERALREIYALPFQ 178
Db    168 SEDPLVSSRTAVAQIKGIQGAGLMTTAKHFAANNQENNRFSVNANVDEQTLREIEFPAFE 227

Qy    179 IAVRDSQPGA--FMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMSDWYGTYSTTEAV 236
Db    228 AS---SKAGAASFMCAYNGLNGKPSCGNDELLNNVLRTQWGFQGWVMSDWLAT-PGTDAI 283

Qy    237 VAGLDLEM----------PGPP-RFRGETLKFNVSNGK-PFIHVIDQRAREVLQFVKK 282
Db    284 TKGLDQEMGVELPGDVPKGEPSPPAKFFGEALKTAVLNGTVPEAAVTRSAERIVGQMEKF 343

Qy    283 CAASGVTENGPETTVNNTPETAALLRKVGNEGIVLLKNENNVLPLS--KKKKTLIVGPNA 340
```

```
Db    344 GLLLATPAPRPE---RDKAGAQAVSRKVAENGAVLLRNEGQALPLAGDAGKSIAVIGPTA 400

Qy    341 KQATYHGGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQCLTPDGAPGM 400

Db    401 VDPKVTGLGSAHVVPDSAAAPLDTIKAR--------AGAGATVTYETGEETFGTQIPAGN 452

Qy    401 RWRVFNEPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELG 460

Db    453 LSPAFNQG------------------HQLE---PGKAGALY---DGTLTVPADGEYRIA 487

Qy    461 LVVCGTAKAYVDDQLVVDNATKQVPGDAFFGSATREETGRINLVKGNTYKFKIEFGSAPT 520

Db    488 VRATG---GYATVQL----GSHTIEAGQVYGKVS---SPLLKLTKG-THKL--------- 527

Qy    521 YTLKGDTIVPGHGSLRVGGCKVIDDQAEIEKSVALAKEHDQVIICAGLNADWETEGADRA 580

Db    528 -TISGFAMSATPLSLELGWVTPAAADATIAKAVESARKARTAVVFA---YDDGTEGVDRP 583

Qy    581 SMKLPGVLDQLIADVAAANPNTVVVMQTGTPEEMPWLDATPAVIQAWYGGNETGNSIADV 640

Db    584 NLSLPGTQDKLISAVADANPNTIVVLNTGSSVLMPWLSKTRAVLDMWYPGQAGAEATAAL 643

Qy    641 VFGDYNPSGKLSLSFPKRLQDNPAFLNFRTEAG----------RTLYGEDVYVGYRYYEF 690

Db    644 LYGDVNPSGKLTQSF-------PAAENQHAVAGDPTSYPGVDNQQTYREGIHVGYRWFDK 696

Qy    691 ADKDVNFPFGHGLSYTTFAFSNLSVSH-KDGKLSVSLSVKNTGSVPGAQVAQLYVKPLQA 749

Db    697 ENVKPLFPFGHGLSYTSFTQSAPTVVRTSTGGLKVTVTVRNSGKRAGQEVVQAYLGASPN 756

Qy    750 AKINRPVKELKGFAKVELQPGETKAVTIEEQEKYVAAYFDEERDQWCVEKGDYEVIVSDS 809

Db    757 VTAPQAKKKLVGYTKVSLAAGEAKTVTVNV----------------------DRRQLQTGS 795

Qy    810 SAAKDGVALRGKFTV 824

Db    796 SSAD----LRGSATV 806


RESULT 5
US-09-147-236-7
; Sequence 7, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
;  APPLICANT: TONOUCHI, Naoto
;  APPLICANT: TSUCHIDA, Takayasu
;  APPLICANT: YOSHINAGA, Fumihiro
;  APPLICANT: TAHARA, Naoki
;  APPLICANT: HAYASHI, Takahisa
;  TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
;  FILE REFERENCE: 6537-011-0PCT
;  CURRENT APPLICATION NUMBER: US/09/147,236A
;  CURRENT FILING DATE: 1999-04-08
;  EARLIER APPLICATION NUMBER: PCT/JP97/03633
;  EARLIER FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 735
;   TYPE: PRT
;   ORGANISM: Acetobacter xylinum
;   FEATURE:
;   OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or
;   OTHER INFORMATION: c
US-09-147-236-7

  Query Match           19.6%;  Score 860;  DB 4;  Length 735;
  Best Local Similarity  28.3%;  Pred. No. 7.9e-76;
  Matches  241;  Conservative  125;  Mismatches  277;  Indels  210;  Gaps   22;

Qy      2 ADIDVEAILKKLTLAEKVDLLAGID----------------FWHTKALPKHGVPSLRFTD 45

Db     37 ADARARQVLASMSLEDKMSLLFSVDGGGFNGSVAPPGGLGSAAYLRAPQGSGLPDLQISD 96

Qy     46 GPNGVRGTKFF--NGVPAACFPCGTSLGSTFNQTLLEEAGKMMGKEAIAKSAHVILGPTI 103

Db     97 AGLGVRNPAHIRRNG-EAVSLPSGQSTASTWDMDMARQAGVMIGREAWQSGFNILLGGGA 155

Qy    104 NMQRSPLGGRGFESIGEDPFLAGLGAAALIRGIQSTGVQATIKHFLCNDQEDRRMMVQSI 163

Db    156 DLTRDPRGGRNFEYAGEDPLQTGRMVGSTIAGVQSQHVISTLKHYAMNDLETSRMTMSAD 215

Qy    164 VTERALREIYALPFQIAVRDSQPGAFMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIM 223

Db    216 IDPVAMRESDLLGFEIALETGHPGAVMCSYNRVNDLYACENPYLLNKTLKQDWHYPGFVM 275

Qy    224 SDWYGTYSTTEAVVAGLDLEMPGP-----PRFRGETLKFNVSNGK-PFIHVIDQRAREVL 277

Db    276 SDWGATHSSARAALAGLDQESAGDHTDARPYFR-TLLAADVKAGRVPEARINDMAER--- 331

Qy    278 QFVKKCAASGVTEN----GPETTVNNTPETAALLRKVGNEGIVLLKNENNVLPLSKKKKT 333

Db    332 -VVRALFAAGLVDHPAQRGPLDVVTDT----LVAQKDEEEGAVLLRNQGNILFLSPTARI 386

Qy    334 LIVGPNAKQATYHGGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQCLT 393

Db    387 AVIGGHADAGVISGGGS----------------------------SQVDPIGGE---- 412

Qy    394 PDGAPGMRWRVFNEPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADE 453

Db    413 ------------------------------------------------------------ 412

Qy    454 DCTYELGLVVCGTAKAYVDDQLVVDNATKQVPGDAFFGSATREETGRINLVKGNTYKFKI 513

Db    413 --------AVKGPGK-------------KEWPGDPVY---------------------- 428

Qy    514 EFGSAPTYTLKGDTIVPGHGSLRVGGCKVIDDQAEIEKSVALAKEHDQVIICAGLNADWE 573

Db    429 -FPSSPLKAMQAE--APG---ARI----TYDPGTSIASAVRAARAADVVVVYA---TQFT 475

Qy    574 TEGADRASMKLPGVLDQLIADVAAANPNTVVVMQTGTPEEMPWLDATPAVIQAWYGGNET 633

Db    476 FEGMDAPSMHLDDNADALITAVAAANPRTVVVMETGDPVLMPWNSSVAGVLEAWFPGSGG 535

Qy    634 GNSIADVVFGDYNPSGKLSLSFPKR------------LQDNPAFLNFRTEAGRTLYGEDV 681

Db    536 GPAIARLLFGKVAPSGHLTMTFPQAESQLAHPDIAGVTADNVFEMQFHTDQ-ELVYDEGS 594

Qy    682 YVGYRYYEFADKDVNFPFGHGLSYTTFAFSNLSVSHKDGKLSVSLSVKNTGSVPGAQVAQ 741

Db    595 DVGYRWFDRNHFKPLYPFGYGLTYTTFSTDGLKVTERHGQVTATFNVHNTGTRAGVDVPQ 654

Qy    742 LYVKPLQAAKINRPVKELKGFAKVELQPGETKAVTIEEQEKYVAAYFDEERDQWCVEKGD 801

Db    655 VYV-----GLPDGGARRLAGWQRISLAPGESRQVSV-QLEPRLLAHFDGKHDRWSVPSGT 708

Qy    802 YEVIVSDSSAAKD 814

Db    709 FRVWL--ASCATD 719


RESULT 6
US-08-462-080B-2
; Sequence 2, Application US/08462080B
; Patent No. 5997913
;  GENERAL INFORMATION:
;    APPLICANT:  Fowler, Timothy
;    APPLICANT:  Barnett, Christopher C.
;    APPLICANT:  Shoemaker, Sharon
;    TITLE OF INVENTION:  Saccharification of Cellulose by Cloning and
;    TITLE OF INVENTION:  Amplification of the Beta-glucosidase Gene of Trichoderma R
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International, Inc.
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
```

```
;      STATE:  Ca
;      COUNTRY:  U.S.A.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,080B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/248,586
;      FILING DATE:  24-MAY-1994
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/807,028
;      FILING DATE:  10-DEC-1991
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/625,140
;      FILING DATE:  10-DEC-1990
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stone, Christopher L.
;      REGISTRATION NUMBER:  35,696
;      REFERENCE/DOCKET NUMBER:  GC78D3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-846-7555
;      TELEFAX:  650-845-6504
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  744 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-462-080B-2

  Query Match             16.7%;  Score 731.5;  DB 2;  Length 744;
  Best Local Similarity   27.5%;  Pred. No. 4.4e-63;
  Matches  234;  Conservative  123;  Mismatches  282;  Indels  211;  Gaps   31;

Qy       7 EAILKKLTLAEKVDLLAGIDFW-------HTKALPKHGVPSLRFTDGPNGVR----GTKF 55
           :| | || | :|| :::|:  |       :|    |   |||   ||| |||     | |
Db      48 KAALAKLNLQDKVGIVSGVG-WNGGPCVGNTSPASKISYPSLCLQDGPLGVRYSTGSTAF 106

Qy      56 FNGVPAACFPCGTSLGSTFNQTLLEEAGKMMGKEAIAKSAHVILGPTIN-MQRSPLGGRG 114
             || ||         ||::  |: | |: :|:|  |   ||||||    : ::| |||
Db     107 TPGVQAA---------STWDVNLIRERGQFIGEEVKASGIHVILGPVAGPLGKTPQGGRN 157

Qy     115 FESIGEDPFLAGLGAAALIRGIQSTGVQATIKHFLCNDQEDRRMMVQSIVTERALREIYA 174
           :|  | ||:| |:     | |||| ||||| ||:: |:||  |  : |   :| | |:|
Db     158 WEGFGVDPYLTGIAMGQTINGIQSVGVQATAKHYILNEQELNRETISSNPDDRTLHELYT 217

Qy     175 LPFQIAVRDSQPGAFMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMSDWYGTYSTTE 234
             ||  ||: :   : | :|| :|     |:   |  :|: : |: | :|:||   ::| :
Db     218 WPFADAVQ-ANVASVMCSYNKVNTTWACEDQYTLQTVLKDQLGFPGYVMTDWNAQHTTVQ 276

Qy     235 AVVAGLDLEMPGPP-----RFRGETLKFNV-SNGKPFIHVIDQRAREVLQFVKKCAASGV 288
           :  :|||: |||       | | |   | ||  |   | |     | : :     :|
Db     277 SANSGLDMSMPGTDFNGNNRLWGPALTNAVNSNQVPTSRVDDM----VTRILAAWYLTGQ 332

Qy     289 TENG-PETTVNNTPE--TAALLRKVGNEGIVLLKNENNVLPLSKKKKTLIVGPNAKQATY 345
           : | |   ::   :      :| :  :|||||||: |:||| |     :|
Db     333 DQAGYPSFNISRNVQGNHKTNVRAIARDGIVLLKNDANILPLKKPASIAVV--------- 383

Qy     346 HGGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQ-CLTPDGAPGMRWRV 404
             ||||:                    :| :    |   :: |   ||| || |
Db     384 ---GSAAI----------------------IGNHARNSPSCNDKGC--DDGALGMGW-- 413

Qy     405 FNEPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELGLVVC 464
               | | :               : | |    | :||

Db     414 ---GSGAVNYPY---------------FVAPYDAINTRASSQGT---------------- 439

Qy     465 GTAKAYVDDQLVVDNATKQVPGDAFFGSATREETGRINLVKGNTYKFKIEFGSAPTYTLK 524
                   |: : |      |     || | :   |  :  ::
Db     440 ---------QVTLSNTDNTSSG----ASAARGKDVAIVFITADS---------------- 470

Qy     525 GDTIVPGHGSLRVGGCKVIDDQAEIEKSVALAKEHDQVIICAGLNADWETEGADRASMKL 584
                 | | : | |     |:  ::                        |    |
Db     471 ------GEGYITVEG--NAGDRNNLD-------------------PWHNGNA------- 495

Qy     585 PGVLDQLIADVAAANPNTVVVMQTGTP---EEMPWLDATPAVIQAWYGGNETGNSIADVV 641
                 |:  || || | :||: :      |::  |    ||: |     |:||:: ||:
Db     496 ------LVQAVAGANSNVIVVVHSVGAIILEQILALPQVKAVVWAGLPSQESGNALVDVL 549

Qy     642 FGDYNPSGKLSLSFPKRLQDNPAFLNFR-TEAGRTLYGEDVYVGYRYYEFADKDVNFPFG 700
           :|| :|||||  :  |    :|   | |    |   : | ::: |:::: |:    : ||
Db     550 WGDVSPSGKLVYTIAK----SPNDYNTRIVSGGSDSFSEGLFIDYKHFDDANITPRYEFG 605

Qy     701 HGLSYTTFAFSNLSV--SHKDGK----------------LSVSLSVKNTGSVPGAQVAQ 741
           :|||| | :| |||  : | |                  :|:: : |:| | ||:|||
Db     606 YGLSYTKFNYSRLSVLSTAKSGPATGAVVPGGPSDLFQNVATVTVDIANSGQVTGAEVAQ 665

Qy     742 LYVK-PLQAAKINRPVKELKGFAKVELQPGETKAVTIEEQEKYVAAYFDEERDQWCVEKG 800
           ||:  |  | :   | |:|:||||: | ||::   |   : : : :|:|   :| | |
Db     666 LYITYPSSAPR--TPPKQLRGFAKLNLTPGQSGTATFNIRRRDL-SYWDTASQKWVVPSG 722

Qy     801 DYEVIVSDSS 810
             : : |  ||
Db     723 SFGISVGASS 732


RESULT 7
US-08-462-090-2
; Sequence 2, Application US/08462090
; Patent No. 6022725
;  GENERAL INFORMATION:
;    APPLICANT:  Fowler, Timothy
;    APPLICANT:  Barnett, Christopher C.
;    APPLICANT:  Shoemaker, Sharon
;    TITLE OF INVENTION:  Saccharification of Cellulose by Cloning
;    TITLE OF INVENTION:  and Amplification of the Beta-glucosidase Gene of
;    TITLE OF INVENTION:  Trichoderma Reesei
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  George Mason Building, 699 Prince St.
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,090
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/625,140
;      FILING DATE:  10-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dillahunty, T. Gene
;      REGISTRATION NUMBER:  25,423
;      REFERENCE/DOCKET NUMBER:  010055-056
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-854-7400
;      TELEFAX:  415-854-8275
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  744 amino acids
;      TYPE:  amino acid
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-462-090-2

  Query Match             16.7%;  Score 731.5;  DB 3;  Length 744;
  Best Local Similarity   27.5%;  Pred. No. 4.4e-63;
  Matches  234;  Conservative  123;  Mismatches  282;  Indels  211;  Gaps   31;

Qy       7 EAILKKLTLAEKVDLLAGIDFW-------HTKALPKHGVPSLRFTDGPNGVR----GTKF 55
           :| | || | :|| :::|:  |       :|    |   |||   ||| |||     | |
Db      48 KAALAKLNLQDKVGIVSGVG-WNGGPCVGNTSPASKISYPSLCLQDGPLGVRYSTGSTAF 106

Qy      56 FNGVPAACFPCGTSLGSTFNQTLLEEAGKMMGKEAIAKSAHVILGPTIN-MQRSPLGGRG 114
             || ||         ||::  |: | |: :|:|  |   ||||||    : ::| |||
Db     107 TPGVQAA---------STWDVNLIRERGQFIGEEVKASGIHVILGPVAGPLGKTPQGGRN 157

Qy     115 FESIGEDPFLAGLGAAALIRGIQSTGVQATIKHFLCNDQEDRRMMVQSIVTERALREIYA 174
           :|  | ||:| |:     | |||| ||||| ||:: |:||  |  : |   :| | |:|
Db     158 WEGFGVDPYLTGIAMGQTINGIQSVGVQATAKHYILNEQELNRETISSNPDDRTLHELYT 217

Qy     175 LPFQIAVRDSQPGAFMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMSDWYGTYSTTE 234
           ||  ||: :   : | :|| :|     |:   |  :|: : |: | :|:||   ::| :
Db     218 WPFADAVQ-ANVASVMCSYNKVNTTWACEDQYTLQTVLKDQLGFPGYVMTDWNAQHTTVQ 276

Qy     235 AVVAGLDLEMPGPP-----RFRGETLKFNV-SNGKPFIHVIDQRAREVLQFVKKCAASGV 288
           :  :|||: |||       | | |   | || |   | |     | : :     :|
Db     277 SANSGLDMSMPGTDFNGNNRLWGPALTNAVNSNQVPTSRVDDM----VTRILAAWYLTGQ 332

Qy     289 TENG-PETTVNNTPE--TAALLRKVGNEGIVLLKNENNVLPLSKKKKTLIVGPNAKQATY 345
           : | |   ::   :      :| :  :|||||||: |:||| |     :|
Db     333 DQAGYPSFNISRNVQGNHKTNVRAIARDGIVLLKNDANILPLKKPASIAVV--------- 383

Qy     346 HGGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQ-CLTPDGAPGMRWRV 404
              ||||:                  :| :    |   :: |   ||| || |
Db     384 ---GSAAI----------------------IGNHARNSPSCNDKGC--DDGALGMGW-- 413

Qy     405 FNEPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELGLVVC 464
           |  |  :                 :  |  |    |  :||
Db     414 ---GSGAVNYPY--------------FVAPYDAINTRASSQGT--------------- 439

Qy     465 GTAKAYVDDQLVVDNATKQVPGDAFFGSATREETGRINLVKGNTYKFKIEFGSAPTYTLK 524
                 |: : |      |     || | :   |  :  ::
Db     440 ---------QVTLSNTDNTSSG----ASAARGKDVAIVFITADS---------------- 470

Qy     525 GDTIVPGHGSLRVGGCKVIDDQAEIEKSVALAKEHDQVIICAGLNADWETEGADRASMKL 584
           |  | : | |     |:  ::                    |    |
Db     471 ------GEGYITVEG--NAGDRNNLD--------------------PWHNGNA------- 495

Qy     585 PGVLDQLIADVAAANPNTVVVMQTGTP---EEMPWLDATPAVIQAWYGGNETGNSIADVV 641
                 |:  || || | :||: :      |::  |    ||: |    |:||:: ||:
Db     496 ------LVQAVAGANSNVIVVVHSVGAIILEQILALPQVKAVVWAGLPSQESGNALVDVL 549

Qy     642 FGDYNPSGKLSLSFPKRLQDNPAFLNFR-TEAGRTLYGEDVYVGYRYYEFADKDVNFPFG 700
           :|| :|||||  :  |    :|   | |    |  : | ::: |:::: |:    : ||
Db     550 WGDVSPSGKLVYTIAK----SPNDYNTRIVSGGSDSFSEGLFIDYKHFDDANITPRYEFG 605

Qy     701 HGLSYTTFAFSNLSV--SHKDGK----------------LSVSLSVKNTGSVPGAQVAQ 741
           :|||||  | :| |||  : | |                 :|:: : |:| | ||:|||
Db     606 YGLSYTKFNYSRLSVLSTAKSGPATGAVVPGGPSDLFQNVATVTVDIANSGQVTGAEVAQ 665

Qy     742 LYVK-PLQAAKINRPVKELKGFAKVELQPGETKAVTIEEQEKYVAAYFDEERDQWCVEKG 800
           ||:  |  | :   | |:|:||||: | ||::   |  : : : :|:|    :| |  |
Db     666 LYITYPSSAPR--TPPKQLRGFAKLNLTPGQSGTATFNIRRRDL-SYWDTASQKWVVPSG 722

Qy     801 DYEVIVSDSS 810
             : :  |  ||
Db     723 SFGISVGASS 732


RESULT 8
US-08-463-461-2
; Sequence 2, Application US/08463461
; Patent No. 6103464
;  GENERAL INFORMATION:
;    APPLICANT:  Fowler, Timothy
;    APPLICANT:  Barnett, Christopher C.
;    APPLICANT:  Shoemaker, Sharon
;    TITLE OF INVENTION:  Saccharification of Cellulose by Cloning
;    TITLE OF INVENTION:  and Amplification of the Beta-glucosidase Gene of
;    TITLE OF INVENTION:  Trichoderma Reesei
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International, Inc.
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94304-1013
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,461
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Christopher L. Stone
;      REGISTRATION NUMBER:  35,696
;      REFERENCE/DOCKET NUMBER:  GC78D4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-846-7555
;      TELEFAX:  415-845-6504
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  744 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-463-461-2

  Query Match             16.7%;  Score 731.5;  DB 3;  Length 744;
  Best Local Similarity   27.5%;  Pred. No. 4.4e-63;
  Matches  234;  Conservative  123;  Mismatches  282;  Indels  211;  Gaps   31;

Qy       7 EAILKKLTLAEKVDLLAGIDFW-------HTKALPKHGVPSLRFTDGPNGVR----GTKF 55
           :| | || | :|| :::|:  |       :|    |   |||   ||| |||     | |
Db      48 KAALAKLNLQDKVGIVSGVG-WNGGPCVGNTSPASKISYPSLCLQDGPLGVRYSTGSTAF 106

Qy      56 FNGVPAACFPCGTSLGSTFNQTLLEEAGKMMGKEAIAKSAHVILGPTIN-MQRSPLGGRG 114
             || ||         ||::  |: | |: :|:|  |   ||||||    : ::| |||
Db     107 TPGVQAA---------STWDVNLIRERGQFIGEEVKASGIHVILGPVAGPLGKTPQGGRN 157

Qy     115 FESIGEDPFLAGLGAAALIRGIQSTGVQATIKHFLCNDQEDRRMMVQSIVTERALREIYA 174
           :|  | ||:| |:     | |||| ||||| ||:: |:||  |  : |   :| | |:|
Db     158 WEGFGVDPYLTGIAMGQTINGIQSVGVQATAKHYILNEQELNRETISSNPDDRTLHELYT 217

Qy     175 LPFQIAVRDSQPGAFMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMSDWYGTYSTTE 234
           ||  ||: :   : | :|| :|     |:   |  :|: : |: | :|:||   ::| :
Db     218 WPFADAVQ-ANVASVMCSYNKVNTTWACEDQYTLQTVLKDQLGFPGYVMTDWNAQHTTVQ 276

Qy     235 AVVAGLDLEMPGPP-----RFRGETLKFNV-SNGKPFIHVIDQRAREVLQFVKKCAASGV 288
           :  :|||: |||       | | |   | || |   | |     | : :     :|
Db     277 SANSGLDMSMPGTDFNGNNRLWGPALTNAVNSNQVPTSRVDDM----VTRILAAWYLTGQ 332

Qy     289 TENG-PETTVNNTPE--TAALLRKVGNEGIVLLKNENNVLPLSKKKKTLIVGPNAKQATY 345
           : | |   ::   :      :| :  :|||||||: |:||| |     :|
Db     333 DQAGYPSFNISRNVQGNHKTNVRAIARDGIVLLKNDANILPLKKPASIAVV--------- 383

Qy     346 HGGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQ-CLTPDGAPGMRWRV 404
              ||||:                  :| :    |   :: |   ||| || |
Db     384 ---GSAAI----------------------IGNHARNSPSCNDKGC--DDGALGMGW-- 413

Qy     405 FNEPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELGLVVC 464
           |  |  :                 :  |  |    |  :||
```

```
Db    414 ---GSGAVNYPY---------------FVAPYDAINTRASSQGT---------------- 439

Qy    465 GTAKAYVDDQLVVDNATKQVPGDAFFGSATREETGRINLVKGNTYKFKIEFGSAPTYTLK 524
                   |: : |     |      || |  :   |   :   ::
Db    440 ---------QVTLSNTDNTSSG----ASAARGKDVAIVFITADS---------------- 470

Qy    525 GDTIVPGHGSLRVGGCKVIDDQAEIEKSVALAKEHDQVIICAGLNADWETEGADRASMKL 584
                | |  : |  |     |:  ::                     |     |
Db    471 ------GEGYITVEG--NAGDRNNLD--------------------PWHNGNA------- 495

Qy    585 PGVLDQLIADVAAANPNTVVVMQTGTP---EEMPWLDATPAVIQAWYGGNETGNSIADVV 641
                |:  || || |  :||: :       |::  |      ||: |       |:||:: ||:
Db    496 ------LVQAVAGANSNVIVVVHSVGAIILEQILALPQVKAVVWAGLPSQESGNALVDVL 549

Qy    642 FGDYNPSGKLSLSFPKRLQDNPAFLNFR-TEAGRTLYGEDVYVGYRYYEFADKDVNFPFG 700
          :|| :|||||  :   |     :|    | |    |     : |  ::: |:::: |:    : ||
Db    550 WGDVSPSGKLVYTIAK----SPNDYNTRIVSGGSDSFSEGLFIDYKHFDDANITPRYEFG 605

Qy    701 HGLSYTTFAFSNLSV--SHKDGK-----------------LSVSLSVKNTGSVPGAQVAQ 741
          :||||| | :| |||   : | |                   :|:: :  |:| | ||:|||
Db    606 YGLSYTKFNYSRLSVLSTAKSGPATGAVVPGGPSDLFQNVATVTVDIANSGQVTGAEVAQ 665

Qy    742 LYVK-PLQAAKINRPVKELKGFAKVELQPGETKAVTIEEQEKYVAAYFDEERDQWCVEKG 800
          ||:  |    |  :    | |:|:||||: | ||::      :|::  :  :  :|:|     :| | |
Db    666 LYITYPSSAPR--TPPKQLRGFAKLNLTPGQSGTATFNIRRRDL-SYWDTASQKWVVPSG 722

Qy    801 DYEVIVSDSS 810
           : :  |   ||
Db    723 SFGISVGASS 732
```

RESULT 9
US-08-981-446B-3
; Sequence 3, Application US/08981446B
; Patent No. 6300112
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: No. 6300112el beta-xylosidase, nucleotide sequence
; TITLE OF INVENTION: encoding it, and use thereof
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,446B
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 804 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger (CBS 120.49)
; STRAIN: NW147
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 1..26
US-08-981-446B-3

```
  Query Match             9.9%;  Score 433.5;  DB 4;  Length 804;
  Best Local Similarity   22.2%;  Pred. No. 1.7e-33;
  Matches  200;  Conservative  131;  Mismatches  276;  Indels  293;  Gaps   36;

Qy     22 LAGIDFWHTKALPKHGVPSLRFTDGPNGVRGTKFFNGVPAACFPCGTSLGSTFNQTLLEE 81
          |   | ::||  ||:      |:|          :|    |   ||       :  |:||: :
Db    107 LPAYQVW-SEAL--HGLDRANFSD-------SGAYNW--ATSFPQPILTTAALNRTLIHQ 154

Qy     82 AGKMMGKEAIA-----KSAHVILGPTINMQRSPLGGRGFESIGEDPFLAGLGAAALIRGI 136
                :: :  |        :      :  | ||   | |: ||| |: |||   || : |     | ||

Db    155 IASIISTQGRAFNNAGRYGLDVYAPNINTFRHPVWGRGQETPGEDVSLAAVYAYEYITGI 214

Qy    137 Q------STGVQATIKHFLCNDQED----RRMMVQSIVTERALREIYALPFQIAVRDSQP 186
          |       :  : || ||:    | |:       |:        :|::  | | |    | :| ||::
Db    215 QGPDPESNLKLAATAKHYAGYDIENWHNHSRLGNDMNITQQDLSEYYTPQFHVAARDAKV 274

Qy    187 GAFMTAYNGINGVSCSENPKYLDGMLRKEWGW--DGLIMSDWYGTY-------------- 230
           : | ||| :|||        :   :|   :||   :|:    | : ||      |
Db    275 QSVMCAYNAVNGVPACADSYFLQTLLRDTFGFVDHGYVSSDCDAAYNIYNPHGYASSQAA 334

Qy    231 STTEAVVAGLDLEMPGPPRFRGETLKFNVSNGKPFIHVIDQRAREVLQFVKKCAASGVTE 290
           :  ||::|| |::           | |          |      :: | | :     : |
Db    335 AAAEAILAGTDIDC-------GTTYQWHLNES---IAAGDLSRDDIEQGVIRLYTTLVQA 384

Qy    291 N--GPETTVNNTP----------ETAA--LLRKVGNEGIVLLKNENNVLPLSKKKKTLIV 336
                       || | |     :    :  :||||||  ||||||:::|
Db    385 GYFDSNTTKANNPYRDLSWSDVLETDAWNISYQAATQGIVLLKNSNNVLPLTEK------ 438

Qy    337 GPNAKQATYHGGGSAALRAYYAVTPFDGLSKQLETPPSYTVGAYTTVPPILGEQCLTPDG 396
                             ||                  ||| |  |          ::|
Db    439 -----------------AY--------------PPSNTTVA------LIGP------- 452

Qy    397 APGMRWRVFNEPPGTPNRQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCT 456
              |
Db    453 -----WA----------------------------------------------------- 454

Qy    457 YELGLVVCGTAKAYVDDQLVVDNATKQVPGDAFFGS--------ATREETG-RINLVKGN 507
                                ||| |: |: ::|:               |  || | ::| :|
Db    455 ----------------------NATTQLLGN-YYGNAPYMISPRAAFEEAGYKVNFAEGT 491

Qy    508 TYKFKIEFGSAPTYTLKGDTIVPGHGSLRVGGCKVIDDQAEIEKSVALAKEHDQVIICAG 567
                 | : | |                         :     ::: |:  |  :|     |
Db    492 --------GISSTST----------------------SGFAAALSAAQSADVIIYAGG 519

Qy    568 LNADWETEGADRASMKLPGVLDQLIADVA-AANPNTVVVMQTGTPEEMPWLDAT------ 620
          ::  | |  || |: ||   ||  :| ||      ::|:| |   :    :|::
Db    520 IDNTLEAEALDRESIAWPGNQLDLIQKLASAAGKKPLIVLQMGGGQ----VDSSSLKNNT 575

Qy    621 --PAVIQAWYGGNETGNSIADVVFGDYNPSGKL-SLSFP-KRLQDNPAF-LNFRTEAGRT 675
            |::     ||   |  :: |:: |    ||:|:| :  :|    :: ||   :| | |
Db    576 NVSALLWGGYPGQSGGFALRDIITGKKNPAGRLVTTQYPASYAEEFPATDMNLRPE---- 631

Qy    676 LYGEDVYVGYRYYEFADKDVNFPFGHGLSYTTFAFSNLSVSHKDGKLSV----------- 724
            |::    |::|         :  : ||||| |||||  |: :  :  |: ||::
Db    632 --GDNPGQTYKWY---TGEAVYEFGHGLFYTTFAESSSNTTTKEVKLNIQDILSQTHEDL 686

Qy    725 -----------SLSVKNTGSVPGAQVAQLYVKPLQAAKINRPVKELKGFAKV-ELQPGET 772
                          : :::||| :       | ::          |      | | | |: :: |: |||
Db    687 ASITQLPVLNFTANIRNTGKLESDYTAMVFANTSDAGPAPYPKKWLVGWDRLGEVKVGET 746

Qy    773 KAVTIEEQEKYVAAYFDEERDQWCVEKGDYEVIVSDSSAAKDGVALRGKFTVGETYWWSG 832
          : : :  :     |      |   |  |   | |:|: ::        : | | |:   |     | |
Db    747 RELRVPVEVGSFARV--NEDGDWVVFPGTFELALNLERKVRVKVVLEGEEEV--VLKWPG 802
```

RESULT 10
US-08-386-727-6
; Sequence 6, Application US/08386727
; Patent No. 5792647
; GENERAL INFORMATION:
; APPLICANT: ROSEMAN, SAUL
; APPLICANT: BASSLER, BONNIE
; APPLICANT: KEYHANI, NEMAT O.
; APPLICANT: CHITLARU, EDITH
; APPLICANT: ROWE, CHRIS
; APPLICANT: YU, CHARLES
; TITLE OF INVENTION: BACTERIAL CATABOLISM OF CHITIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.

```
;       COUNTRY:  USA
;       ZIP:  20005
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/386,727
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  HOBBS, ANN S.
;       REGISTRATION NUMBER:  36,830
;       REFERENCE/DOCKET NUMBER:  4130/206916
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-861-3000
;       TELEFAX:  202-822-0944
;       TELEX:  6714627 CUSH
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  328 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-386-727-6

  Query Match             3.0%;  Score 132.5;  DB 1;  Length 328;
  Best Local Similarity   23.4%;  Pred. No. 0.00022;
  Matches   62;  Conservative   37;  Mismatches  103;  Indels   63;  Gaps    9;

Qy      65 PCGTSLGSTFNQT-LLEEAGKMMGKEAIAKSAHVILGPTIN--MQRSPLGGRGFESIGED 121
           ||      : | | | |: | :|  | ||    :  | ::       :| | |   |:|
Db      78 PCAQLYARSDNGTQLAEDGGWLMAAELIAHDIDLSFAPVLDKGFDCRAIGNRAF---GDD 134

Qy     122 PFLAGLGAAALIRGIQSTGVQATIKHF----------LCNDQEDRRMMVQSIVTERALR 170
              ::| :||::| |:  | |||          |    ::|  :   :   ||
Db     135 VQTVLTYSSAYMRGMKSVGMATTGKHFPGHGAVIADSHLETPYDERDSIADDMTIFRAQI 194

Qy     171 EIYALPFQIAVR------DSQPGAFMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMS 224
           |   |   :         |:||              | :| :|  :||:| |: |:: |
Db     195 EAGILDAMMPAHVIYPHYDAQP-------------ASGSPYWLKQVLRQELGFQGIVFS 240

Qy     225 DWYGTYSTTEAVVAGLDLEMPGPPRFRGETLKFNVSNGKPFIHVIDQRAREVLQFVKKCA 284
           |               || | |     |   :  |       :|    ||   |: :
Db     241 D--------------DLSMEGAAIMGGPAERAQQS--------LDAGCDMVLMCNKRES 277

Qy     285 ASGVTENGPETTVNNTPETAALLRK 309
           |  | :  | : |   |:  :||::
Db     278 AVAVLDQLPISVV---PQAQSLLKQ 299


RESULT 11
US-08-600-452A-6
; Sequence 6, Application US/08600452A
; Patent No. 5985644
;  GENERAL INFORMATION:
;    APPLICANT:  ROSEMAN, SAUL
;    APPLICANT:  BASSLER, BONNIE
;    APPLICANT:  KEYHANI, NEMAT O.
;    APPLICANT:  CHITLARU, EDITH
;    APPLICANT:  ROWE, CHRIS
;    APPLICANT:  YU, CHARLES
;    TITLE OF INVENTION:  BACTERIAL CATABOLISM OF CHITIN
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FISH & RICHARDSON P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,452A
;      FILING DATE:  13-FEB-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07662/005001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 678-5070
;      TELEFAX:  (619) 678-5099
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  328 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-600-452A-6

  Query Match             3.0%;  Score 132.5;  DB 2;  Length 328;
  Best Local Similarity   23.4%;  Pred. No. 0.00022;
  Matches   62;  Conservative   37;  Mismatches  103;  Indels   63;  Gaps    9;

Qy      65 PCGTSLGSTFNQT-LLEEAGKMMGKEAIAKSAHVILGPTIN--MQRSPLGGRGFESIGED 121
           ||      : | | | |: | :|  | ||    :  | ::       :| | |   |:|
Db      78 PCAQLYARSDNGTQLAEDGGWLMAAELIAHDIDLSFAPVLDKGFDCRAIGNRAF---GDD 134

Qy     122 PFLAGLGAAALIRGIQSTGVQATIKHF----------LCNDQEDRRMMVQSIVTERALR 170
              ::| :||::| |:  | |||          |    ::|  :   :   ||
Db     135 VQTVLTYSSAYMRGMKSVGMATTGKHFPGHGAVIADSHLETPYDERDSIADDMTIFRAQI 194

Qy     171 EIYALPFQIAVR------DSQPGAFMTAYNGINGVSCSENPKYLDGMLRKEWGWDGLIMS 224
           |   |   :         |:||              | :| :|  :||:| |: |:: |
Db     195 EAGILDAMMPAHVIYPHYDAQP-------------ASGSPYWLKQVLRQELGFQGIVFS 240

Qy     225 DWYGTYSTTEAVVAGLDLEMPGPPRFRGETLKFNVSNGKPFIHVIDQRAREVLQFVKKCA 284
           |               || | |     |   :  |       :|    ||   |: :
Db     241 D--------------DLSMEGAAIMGGPAERAQQS--------LDAGCDMVLMCNKRES 277

Qy     285 ASGVTENGPETTVNNTPETAALLRK 309
           |  | :  | : |   |:  :||::
Db     278 AVAVLDQLPISVV---PQAQSLLKQ 299


RESULT 12
US-09-071-035-458
; Sequence 458, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
```

```
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  458:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2032 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-071-035-458

  Query Match             2.9%;  Score 128.5;  DB 4;  Length 2032;
  Best Local Similarity   20.0%;  Pred. No. 0.015;
  Matches  126;  Conservative   81;  Mismatches  227;  Indels  197;  Gaps   31;

Qy    227 YGTYS--------------TTEAVVAG---LDLEMPGPPRFRGETLKFNVSNGK-PFIH 267
          ||||:               |:|: : |   ||  :              | |:|: |
Db    295 YGTYTISEDGTVRFTFNERITSESDIHGDFSLDTHL------------NDSDGRGPGDW 341

Qy    268 VIDQRAREVLQFVKKCAASGVTENGPETTVNNTPETAALLRKVGN-------EGIVLLKN 320
          |||   :| |                |   :   |:|   : | |:          :
Db    342 VIDIPTQEDL---------------PPVVIPIVPDTEQQIDKQGHFDRTPNPSAITWTVD 386

Qy    321 ENNVLPLSKKKKTLIVGPNAKQATYHGGGSAALRAYYAVTPFDGLSKQLE---TPPSYTV 377
           |  :   | :|   |   :   |    :::  |  |    ||  |::    :|  |||
Db    387 INQAM----KDQT---NPTVTETWPTGNTFKSVKVYELVMNLDGTIKEVGRELSPDEYTV 439

Qy    378 G------------AY-----TTVPPILGEQCLTPDGAPGMRWRVF------NEPPGTPN 413
                       ||     ||:    :: : |||   : ::        | | |
Db    440 DKNGNVTIKGDTNKAYRLEYQTTI-----DEAVIPDGGGDVPFKNHATLTSDNNPNGLDA 494

Qy    414 RQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELGLVVCGTAKAYVDD 473
            :   :     |   :||     ||                :  |  |      :| : |
Db    495 EATVTATYGKMLDKRNIDYDEANQEFTW-----------EINYNYGEQTIPKDQAVITD 542

Qy    474 QLVVDNATKQVPGDAFFGSATREETGR----INLVKGNTYKFKIEFGSAPTYTLKGDTIV 529
           :  || | :  |                        ||:|  ||                ::
Db    543 TM-GDNLTFE-PDSLHLYSVTFDDKGNEVVGAELVEGKDYK---------------VVI 584

Qy    530 PGHGSLR------VGGCKVIDDQAE----IEKSVALAKEHDQVIICAGLNADWETEGADR 579
           | ||              | |    || : :    :|  ||:    :::|  :  | : :  : :
Db    585 NGDGSFAIDFLHDVTGAVKIDYKTKVDGIVEGDVAV---NNRVDVGTGQHSEDDGTASQQ 641

Qy    580 ASMKLPGVLD---QLIADVAAANPNTVVV---MQTGTPEEMPWLDATPAVIQAWYGGNET 633
           :|  | :|      |      | | |  ::    : | | | :| |    |
Db    642 NIIKNTGAVDYQNSTIGWTLAVNQNNYLMENAVITDTYEPVPGLTMVP-----------: 689

Qy    634 GNSIADVVFGDYNPSGKLSLSFPKRLQDNPAFLNFRTEAGRTLYGEDVYVGYR------Y 687
           ||:  |  |     |:|:                 :|  |  |  ||  |::          |
Db    690 -NSL---VVKDTTTGAQLTLG-----------KDFMVEITRNADGE---TGFKVSFIGAY 731

Qy    688 YEFADKDVNFPFGHGLSYTTFAFSNLSVSHKDGKLSVSLSVKNTGSVPGAQVAQLYVKPL 747
           : :|           ::||||       |:  |            :  ::   |
Db    732 AKTSD-------AFHITYTTF----FDVTELDANNPALDHYRNTAAIDWTDEA----GNN 776

Qy    748 QAAKINRPVKELKGFAKVELQPGETKAVTIE 778
           :: ::| |  |  |       :  |    ||| |
Db    777 HHSEDSKPFKPLPAFDLNAQKSGVYNAVTKE 807

RESULT 13
US-09-071-035-462
; Sequence 462, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  462:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2032 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-071-035-462

  Query Match             2.9%;  Score 128.5;  DB 4;  Length 2032;
  Best Local Similarity   20.0%;  Pred. No. 0.015;
  Matches  126;  Conservative   81;  Mismatches  227;  Indels  197;  Gaps   31;

Qy    227 YGTYS--------------TTEAVVAG---LDLEMPGPPRFRGETLKFNVSNGK-PFIH 267
          ||||:               |:|: : |   ||  :              | |:|: |
Db    295 YGTYTISEDGTVRFTFNERITSESDIHGDFSLDTHL------------NDSDGRGPGDW 341

Qy    268 VIDQRAREVLQFVKKCAASGVTENGPETTVNNTPETAALLRKVGN-------EGIVLLKN 320
          |||   :| |                |   :   |:|   : | |:          :
Db    342 VIDIPTQEDL---------------PPVVIPIVPDTEQQIDKQGHFDRTPNPSAITWTVD 386

Qy    321 ENNVLPLSKKKKTLIVGPNAKQATYHGGGSAALRAYYAVTPFDGLSKQLE---TPPSYTV 377
           |  :   | :|   |   :   |    :::  |  |    ||  |::    :|  |||
Db    387 INQAM----KDQT---NPTVTETWPTGNTFKSVKVYELVMNLDGTIKEVGRELSPDEYTV 439

Qy    378 G------------AY-----TTVPPILGEQCLTPDGAPGMRWRVF------NEPPGTPN 413
                       ||     ||:    :: : |||   : ::        | | |
Db    440 DKNGNVTIKGDTNKAYRLEYQTTI-----DEAVIPDGGGDVPFKNHATLTSDNNPNGLDA 494

Qy    414 RQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELGLVVCGTAKAYVDD 473
            :   :     |   :||     ||                :  |  |      :| : |
Db    495 EATVTATYGKMLDKRNIDYDEANQEFTW-----------EINYNYGEQTIPKDQAVITD 542

Qy    474 QLVVDNATKQVPGDAFFGSATREETGR----INLVKGNTYKFKIEFGSAPTYTLKGDTIV 529
           :  || | :  |                        ||:|  ||                ::
Db    543 TM-GDNLTFE-PDSLHLYSVTFDDKGNEVVGAELVEGKDYK---------------VVI 584

Qy    530 PGHGSLR------VGGCKVIDDQAE----IEKSVALAKEHDQVIICAGLNADWETEGADR 579
           | ||              | |    || : :    :|  ||:    :::|  :  | : :  : :
Db    585 NGDGSFAIDFLHDVTGAVKIDYKTKVDGIVEGDVAV---NNRVDVGTGQHSEDDGTASQQ 641

Qy    580 ASMKLPGVLD---QLIADVAAANPNTVVV---MQTGTPEEMPWLDATPAVIQAWYGGNET 633
```

```
Db    642 NIIKNTGAVDYQNSTIGWTLAVNQNNYLMENAVITDTYEPVPGLTMVP---------- 689

Qy    634 GNSIADVVFGDYNPSGKLSLSFPKRLQDNPAFLNFRTEAGRTLYGEDVYVGYR------Y 687
Db    690 -NSL---VVKDTTTGAQLTLG----------KDFMVEITRNADGE---TGFKVSFIGAY 731

Qy    688 YEFADKDVNFPFGHGLSYTTFAFSNLSVSHKDGKLSVSLSVKNTGSVPGAQVAQLYVKPL 747
Db    732 AKTSD-------AFHITYTTF----FDVTELDANNPALDHYRNTAAIDWTDEA----GNN 776

Qy    748 QAAKINRPVKELKGFAKVELQPGETKAVTIE 778
Db    777 HHSEDSKPFKPLPAFDLNAQKSGVYNAVTKE 807


RESULT 14
US-09-071-035-466
; Sequence 466, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  466:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2032 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-071-035-466

  Query Match             2.9%;  Score 128.5;  DB 4;  Length 2032;
  Best Local Similarity   20.0%;  Pred. No. 0.015;
  Matches  126;  Conservative   81;  Mismatches  227;  Indels  197;  Gaps   31;

Qy    227 YGTYS--------------TTEAVVAG---LDLEMPGPPRFRGETLKFNVSNGK-PFIH 267
Db    295 YGTYTISEDGTVRFTFNERITSESDIHGDFSLDTHL------------NDSDGRGPGDW 341

Qy    268 VIDQRAREVLQFVKKCAASGVTENGPETTVNNTPETAALLRKVGN-------EGIVLLKN 320
Db    342 VIDIPTQEDL--------------PPVVIPIVPDTEQQIDKQGHFDRTPNPSAITWTVD 386

Qy    321 ENNVLPLSKKKKTLIVGPNAKQATYHGGGSAALRAYYAVTPFDGLSKQLE---TPPSYTV 377
Db    387 INQAM----KDQT---NPTVTETWPTGNTFKSVKVYELVMNLDGTIKEVGRELSPDEYTV 439

Qy    378 G------------AY-----TTVPPILGEQCLTPDGAPGMRWRVF------NEPPGTPN 413
Db    440 DKNGNVTIKGDTNKAYRLEYQTTI-----DEAVIPDGGGDVPFKNHATLTSDNNPNGLDA 494

Qy    414 RQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELGLVVCGTAKAYVDD 473
Db    495 EATVTATYGKMLDKRNIDYDEANQEFTW-----------EINYNYGEQTIPKDQAVITD 542

Qy    474 QLVVDNATKQVPGDAFFGSATREETGR----INLVKGNTYKFKIEFGSAPTYTLKGDTIV 529
Db    543 TM-GDNLTFE-PDSLHLYSVTFDDKGNEVVGAELVEGKDYK---------------VVI 584

Qy    530 PGHGSLR------VGGCKVIDDQAE----IEKSVALAKEHDQVIICAGLNADWETEGADR 579
Db    585 NGDGSFAIDFLHDVTGAVKIDYKTKVDGIVEGDVAV---NNRVDVGTGQHSEDDGTASQQ 641

Qy    580 ASMKLPGVLD---QLIADVAAANPNTVVV---MQTGTPEEMPWLDATPAVIQAWYGGNET 633
Db    642 NIIKNTGAVDYQNSTIGWTLAVNQNNYLMENAVITDTYEPVPGLTMVP---------- 689

Qy    634 GNSIADVVFGDYNPSGKLSLSFPKRLQDNPAFLNFRTEAGRTLYGEDVYVGYR------Y 687
Db    690 -NSL---VVKDTTTGAQLTLG----------KDFMVEITRNADGE---TGFKVSFIGAY 731

Qy    688 YEFADKDVNFPFGHGLSYTTFAFSNLSVSHKDGKLSVSLSVKNTGSVPGAQVAQLYVKPL 747
Db    732 AKTSD-------AFHITYTTF----FDVTELDANNPALDHYRNTAAIDWTDEA----GNN 776

Qy    748 QAAKINRPVKELKGFAKVELQPGETKAVTIE 778
Db    777 HHSEDSKPFKPLPAFDLNAQKSGVYNAVTKE 807


RESULT 15
US-09-071-035-460
; Sequence 460, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  460:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  737 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-071-035-460

  Query Match             2.8%;  Score 122.5;  DB 4;  Length 737;
  Best Local Similarity   19.8%;  Pred. No. 0.0095;
  Matches  116;  Conservative   76;  Mismatches  202;  Indels  193;  Gaps   30;

Qy    227 YGTYS--------------TTEAVVAG---LDLEMPGPPRFRGETLKFNVSNGK-PFIH 267
          ||||:                |:|: : |   ||  :             | |:|: |
Db    263 YGTYTISEDGTVRFTFNERITSESDIHGDFSLDTHL------------NDSDGRGPGDW 309

Qy    268 VIDQRAREVLQFVKKCAASGVTENGPETTVNNTPETAALLRKVGN-------EGIVLLKN 320
          |||   :| |               |   :   |:|   : | |:         |    :
Db    310 VIDIPTQEDL---------------PPVVIPIVPDTEQQIDKQGHFDRTPNPSAITWTVD 354

Qy    321 ENNVLPLSKKKKTLIVGPNAKQATYHGGGSAALRAYYAVTPFDGLSKQLE---TPPSYTV 377
           |  :    | :|    |   :    |    ::: |  |   ||  |::    :|  |||
Db    355 INQAM----KDQT---NPTVTETWPTGNTFKSVKVYELVMNLDGTIKEVGRELSPDEYTV 407

Qy    378 G-------------AY-----TTVPPILGEQCLTPDGAPGMRWRVF------NEPPGTPN 413
                        ||     ||:     :: : |||   : ::        | | |
Db    408 DKNGNVTIKGDTNKAYRLEYQTTI-----DEAVIPDGGGDVPFKNHATLTSDNNPNGLDA 462

Qy    414 RQHIDELFFTKTDMHLVDYYHPKAADTWYADMEGTYTADEDCTYELGLVVCGTAKAYVDD 473
             :   :    |   :||       ||            :  |  |       :| : |
Db    463 EATVTATYGKMLDKRNIDYDEANQEFTW-----------EINYNYGEQTIPKDQAVITD 510

Qy    474 QLVVDNATKQVPGDAFFGSATREETGR----INLVKGNTYKFKIEFGSAPTYTLKGDTIV 529
           :  || | : |       | | :: |       ||:|  ||                 ::
Db    511 TM-GDNLTFE-PDSLHLYSVTFDDKGNEVVGAELVEGKDYK---------------VVI 552

Qy    530 PGHGSLR------VGGCKVIDDQAE----IEKSVALAKEHDQVIICAGLNADWETEGADR 579
           | ||        | |   || : :    :|  ||:   :::| :  | ::: :   : :
Db    553 NGDGSFAIDFLHDVTGAVKIDYKTKVDGIVEGDVAV---NNRVDVGTGQHSEDDGTASQQ 609

Qy    580 ASMKLPGVLD---QLIADVAAANPNTVVV---MQTGTPEEMPWLDATPAVIQAWYGGNET 633
             :|  | :|     |    | | |  ::   : | | | :| |    |
Db    610 NIIKNTGAVDYQNSTIGWTLAVNQNNYLMENAVITDTYEPVPGLTMVP---------- 657

Qy    634 GNSIADVVFGDYNPSGKLSLSFPKRLQDNPAFLNFRTEAGRTLYGEDVYVGYR------Y 687
           ||:   |  |     :|:|             :|  |  |   ||    |::      |
Db    658 -NSL---VVKDTTTGAQLTLG----------KDFMVEITRNADGE---TGFKVSFIGAY 699

Qy    688 YEFADKDVNFPFGHGLSYTTFAFSNLSVSHKDGKLSVSLSVKNTGSV 734
           : :|          ::|||||      |:  |          :|| ::
Db    700 AKTSD-------AFHITYTTF----FDVTELDANNPALDHYRNTAAI 735


Search completed: April 26, 2003, 13:02:52
Job time : 27 secs
```